MEMORANDUM FOR SECRETARIES OF THE MILITARY DEPARTMENTS
CHAIRMAN OF THE JOINT CHIEFS OF STAFF
UNDER SECRETARIES OF DEFENSE
DIRECTOR, DEFENSE RESEARCH AND EVALUATION
ASSISTANT SECRETARIES OF DEFENSE
GENERAL COUNSEL OF THE DEPARTMENT OF DEFENSE
INSPECTOR GENERAL OF THE DEPARTMENT OF DEFENSE
DIRECTOR OPERATIONAL TEST AND EVALUATION
ASSISTANTS TO THE SECRETARY OF DEFENSE
DIRECTOR, ADMINISTRATION AND MANAGEMENT
DIRECTOR, FORCE TRANSFORMATION
DIRECTOR, NET ASSESSMENT
DIRECTORS OF THE DEFENSE AGENCIES
DIRECTORS OF THE DOD FIELD ACTIVITIES

SUBJECT: DoD Net-Centric Data Strategy: Visibility—Advertising Data Assets with Discovery Metadata

Reference: a) Department of Defense Net-Centric Data Strategy, 9 May 2003
b) Department of Defense Discovery Metadata Standard (DDMS), *correct reference here*
c) Global Information Grid Enterprise Services (GIS ES) Implementation, *update reference here*

This memorandum provides guidance on planning for and implementing data asset "visibility" as described in the Net-Centric Data Strategy. In accordance with Reference a) Section 3.1, data producers make their data assets visible by "advertising" their data assets—identifying what the assets are—and making the assets available via shared spaces on the Global Information Grid (GIG). Advertising data assets requires associating descriptive information (hereafter referred to as discovery metadata) with each asset and providing the discovery metadata to an appropriate metadata catalog or repository. Data assets that are advertised provide higher value because a broader base of users can exploit them.

Discovery metadata must be provided for all data assets that will be available on the GIG. Discovery metadata must be generated and catalogued at the time a data asset is made available to the Enterprise or Communities of Interest (COIs). A DoD Discovery Metadata Standard (DDMS), Reference b), has been developed to foster consistency in data asset descriptions.

Further clarification of terminology, implementation considerations, and examples are provided in the Attachment.

Accordingly, the following actions are directed:

The Defense Information Systems Agency (DISA) shall begin planning for and implementing capabilities required to exploit discovery metadata.

1) Not later than December 1, 2004 *(TBR)*, DISA will deploy a tool to register existing or developmental discovery capabilities, such as search engines, metadata catalogs, and content management systems, that identify, catalog, locate and access data assets available on Defense Networks. DISA will establish a simple, user-friendly registration process for these capabilities. DISA is to use this information to assist the Department in planning the transition from existing discovery capabilities to the future GIG Enterprise Services (GIG ES) discovery capability.

2) As of January 1, 2004 *(TBR)*, a draft discovery interface standard consistent with the DDMS and draft requirements for the GIG ES discovery capability will be provided to Components and cognizant OSD authorities for their use in transition planning and implementation in accordance with Reference c).

3) An initial Enterprise Discovery capability available to the Enterprise that leverages DDMS compliant discovery metadata will be provided in the initial deployment of GIG ES currently planned for FY05.

Components and cognizant OSD authorities shall begin planning for and implementing the advertisement of data assets by associating and cataloging discovery metadata per the following:

4) As of October 1, 2003 *(TBR)*, all applications, systems and platforms being developed, procured or acquired will be required to be able to associate DDMS compliant discovery metadata with their expected data assets.

5) As of January 15, 2004 *(TBR)*, existing or developmental discovery capabilities that identify, catalog, locate and access data assets available on Defense Networks will be registered via the process established by DISA. Search engines, metadata catalogs, and content management systems should be registered, but not relational databases or file systems.

6) As of June 1, 2004 *(TBR)*, all existing operational systems and applications that provide metadata for discovery purposes must use DDMS compliant discovery metadata.

7) In accordance with Reference c), each DoD Component and cognizant OSD authority will provide their GIG ES Implementation plans within 90 days after an approved Net-Centric Enterprise Services (NCES) program milestone B. To address the data visibility approaches in Reference a), these plans will include:

   a) Approach and schedule for making systems and applications capable of generating and associating DDMS compliant discovery metadata with data assets available on the GIG

   b) Approach and schedule for producing discovery metadata in response to an Enterprise search query in conformance with the DISA supplied interface standard and Enterprise Discovery service requirements.

   c) Analysis of the impact and value of associating discovery metadata with assets from retired systems, platforms or applications, and plans for addressing those assets that should be advertised.

Subsequent guidance on other aspects of Reference a) including accessibility (posting of data assets to shared spaces) and usability (registering structural metadata for data assets for understandability and interoperability) will be provided. A forthcoming DoD data directive will encompass these guidance memos.

The GIG ES metadata standards working group will adopt and configuration manage the DDMS. Questions about the DDMS should be directed to Mr. Clay Robinson at 703-602-0882 or Clay.Robinson@osd.mil.

Your involvement and assistance are critical to the success of the Department's net-centric transformation. Addressees are requested to assure widest distribution of this memorandum. The Director, Joint Staff is requested to forward this memorandum to the Unified Combatant Commands. My point-of-contact for this action is Mr. Anthony Simon at 703-602-1090 or Anthony.Simon@osd.mil.

John P. Stenbit

ATTACHMENT

Terminology, Implementation Considerations, and Examples

"Data asset" refers to any electronic entity that is composed of or provides access to data. In this memorandum, data asset is further defined as system or application output files and streams, databases, electronic documents, electronic records, and web pages. Additionally, data assets are services offered by a system or application that provide access to that system or application's data. For example, a database is a data asset that contains data records. Similarly, a query service that returns individual records from a database would be a data asset, or a web site that returns data in response to specific user requests (e.g., weather.com) would be a data asset. Data assets are produced by humans using applications (e.g., word processing) or by systems and platforms (e.g., unmanned aerial vehicles). The data within a data asset may be static or dynamic.

Examples of data assets include a document, an electronic file folder that contains multiple documents with similar content, or a folder that contains documents grouped by some other relationship. If a particular file folder or storage location contains many documents that are all very similar (e.g., a folder that holds daily logs files of network activity) then it may be appropriate to only provide discovery metadata for the folder as a whole. If a file folder or storage location contains many documents that are related but not likely to be the subject of a single search (e.g., a folder containing daily intelligence reports), then it is desirable that each document contains discovery metadata rather than simply providing discovery metadata for the folder as a whole.

As another example, a database may be treated as a single asset rather than identifying each individual record as an asset. While individual records are data assets, and can each have associated discovery metadata, the initial thrust of the Department's discovery capability is to provide visibility of the database that contains those individual records (i.e., make the database itself known to potential users). Where possible, the system populating that database asset may provide a mechanism (e.g., web service, web portal) that also allows for individual records to be discovered in accordance with Reference b). Hence, the existence of the database as an asset is made visible to the Enterprise or COIs and discovery metadata for individual records may be provided in response to a query to the database.

Systems that generate real-time, streaming, or dynamic data (e.g., unmanned aerial vehicle data feeds) should provide a mechanism that allows access to the data in a manner that does not negatively impact mission critical processes. For example, a real-time, streaming video system may opt to offload incoming data to a secondary storage location that has a software service that allows users to search and access the recently stored data. In this example, the service that provides access to the offloaded data should be advertised to the Enterprise in accordance with Reference b) and, as in the database example, the service should be capable of generating DDMS compliant discovery metadata for segments of content within that stream.

As demonstrated in the previous examples, identifying data assets and determining the appropriate level for associating discovery metadata is a key step in advertising assets. Discovery metadata is used by systems or applications to search for applicable information and by humans who can view the discovery metadata to determine the relevancy of the associated data. To adhere to the intent of Reference a), Components and cognizant OSD authorities should generate

discovery metadata that allows for Enterprise or Community discovery of assets and supports evaluation of usefulness by unanticipated users. **Components, programs or COIs have discretion in identifying and prioritizing data assets for associating discovery metadata. Components, programs and COIs should use common sense and engineering judgment to determine the level of data assets for which to generate and store discovery metadata.**

Discovery metadata can be associated with a data asset using a variety of methods. For example, an electronic metacard (a record containing the necessary discovery metadata) can be generated and stored separately but linked to the asset; an electronic metacard can be embedded in the asset; or an application can be provided that dynamically generates a metacard for an asset in response to a search query. The discovery metadata that needs to be generated is described in Reference b). Per the DDMS, discovery metadata for an asset can be generated and retained in any format; however, once the Enterprise and COI discovery interface standard is defined, discovery metadata must be provided to search queries in accordance with the interface standard. This model provides flexibility for system developers to generate and retain discovery metadata using an approach that meets system needs. When the Enterprise discovery interface standard is defined and implemented in an Enterprise Discovery service, Components and cognizant OSD authorities should have DDMS compliant discovery metadata available for their data assets. Hence, for Components and cognizant OSD authorities to meet the discovery interface standard will require that their existing discovery metadata be formatted per the interface standard. The interface standard will use the information elements listed in the DDMS so that systems will not be required to generate any additional discovery metadata, or change their approach to providing metadata, in order to participate in Enterprise discovery.

For metacards that are generated and stored, they must be stored, or cataloged, in a location that is, or will be, accessible on the GIG for future use by Enterprise and COI discovery services. Services that provide dynamic generation of metacards in response to a discovery query must also be available on the GIG (e.g., a web-based service).